Introduction to quantum computation: Finding Eigenvalues and The Hidden Subgroup Problem

Lecture Notes for Quantum Computing

Lecturer: Dorit Aharonov
Scribe: Yotam Drier and Eyal David

April 15-19, 2001

Summary of the fifth week of semester 2 (15.4.2001 - 19.4.2001)

1 Phase Estimation: Finding an Eigenvalue

If U is a unitary matrix, and $|e\rangle$ is its eigenvector then $U|e\rangle = \lambda|e\rangle$. U is unitary and therefore $|\lambda| = 1$, and so $\lambda = e^{2\pi i\theta}$. we would like to be able to compute $|e\rangle|0\rangle \rightarrow |e\rangle|\lambda\rangle$ or $|e\rangle|0\rangle \rightarrow |e\rangle|\theta\rangle$.

1.1 First solution (unefficient)

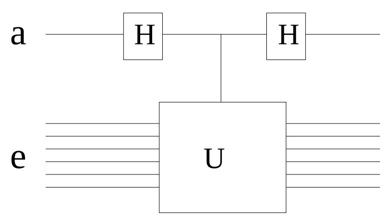

Figure 1: Controled-U

The circuit above performs the following sequence of operations :

$$|0\rangle|e\rangle \rightarrow^{H}$$

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|e\rangle \rightarrow^{controled-U}$$

$$\frac{1}{\sqrt{2}}(|0\rangle|e\rangle + \lambda|1\rangle|e\rangle) \rightarrow^{H}$$

$$\frac{1}{\sqrt{2}}[\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) + \frac{\lambda}{\sqrt{2}}(|0\rangle - |1\rangle)]|e\rangle = \frac{1}{2}[(1+\lambda)|0\rangle + (1-\lambda)|1\rangle]|e\rangle$$

We can now measure the first bit. Since $\lambda = e^{2\pi i\theta}$,

$$Pr(|0\rangle) = \frac{|1 + cos(2\pi\theta) + isin(2\pi\theta)|^2}{4} = \frac{(1 + cos(2\pi\theta))^2 + sin^2(2\pi\theta)}{4} =$$

$$\frac{1 + cos(2\pi\theta)}{2}$$

and in a similar way

$$Pr(|1\rangle) = \frac{1 - cos(2\pi\theta)}{2}$$

Notice that if $\lambda = 1(-1)$ then if we will measure the output of the circuit we will receive $0(1)$ with probability 1, thus we can recover λ. In general we can still get some information about λ, by composing the circuit to itself (as shown below).

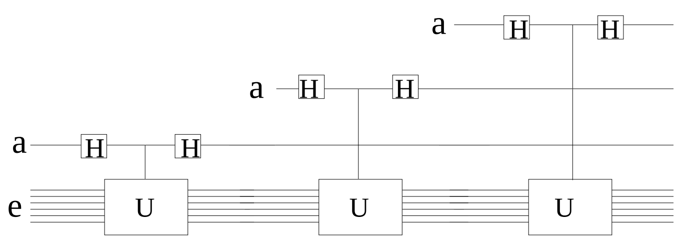

Figure 2: Multiple Controled-U

This circuit enables us to perform a number of independent trials, thereby recovering more information about $\cos\theta$.

If instead of using H (in the first time) we use the transform $|0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ then we will get :

$$|0\rangle|e\rangle \rightarrow^{H'}$$

$$\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)|e\rangle \rightarrow^{controled-U}$$

$$\frac{1}{\sqrt{2}}(|0\rangle|e\rangle + i\lambda|1\rangle|e\rangle) \rightarrow^{H}$$

$$\frac{1}{\sqrt{2}}[\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) + \frac{i\lambda}{\sqrt{2}}(|0\rangle - |1\rangle)]|e\rangle = \frac{1}{2}[(1+i\lambda)|0\rangle + (1-i\lambda)|1\rangle]|e\rangle$$

$$Pr(|0\rangle) = \frac{|1 - sin(2\pi\theta) + icos(2\pi\theta)|^2}{4} = \frac{(1 - sin(2\pi\theta))^2 + cos^2(2\pi\theta)}{4} =$$

$$\frac{1 - sin(2\pi\theta)}{2}$$

and in a similar way

$$Pr(|1\rangle) = \frac{1 + sin(2\pi\theta)}{2}$$

now we can estimate $sin\theta$ as well, and from $sin\theta$ and $cos\theta$ we can easily compute θ. This is the problem of estimating the bias of a coin, within ϵ, with probability $1 - \delta$. As we know this requires $\Theta(\frac{log(\frac{1}{\delta})}{\epsilon^2})$ samples. So, in order to achieve n bits of accuracy we would need approximately 4^n samples.

1.2 A better solution (with additional assumption)

Suppose we can compute U^{2^l} in a polynomial time. This allows us to use a more efficient method. Notice that $U^{2^l}|e\rangle = \lambda^{2^l}|e\rangle = e^{2\pi i 2^l \theta}$. We can use the first algorithm to estimate θ within $\epsilon = \frac{1}{8}$. Now we will repeat the algorithm for $U^2, U^4, ..., U^{2^k}$. Each time we compute $2^k\theta$ and that enables us to estimate θ with greater accuracy using the following lemma :

Lemma 1 *assuming we know $\theta, 2\theta, ..., 2^l\theta$ within $\epsilon = \frac{1}{8}$ we can evaluate θ within $\frac{\epsilon}{2^{l-1}}$*

Proof Given $2^l\theta$ we can extrapolate two possible values of $2^{l-1}\theta$. However, only one such value corresponds with the measurement of $2^{l-1}\theta$, because $2\theta = 2(\theta + \pi)$, and $\epsilon < \pi$. Obviously, knowing $2^l\theta$ within ϵ enables us to

know $2^{l-1}\theta$ within $\frac{\epsilon}{2}$. We can continue this way recursively, and find θ within $\frac{\epsilon}{2^{l-1}}$. *QED.*
The algorithm requires k measurements to achieve k bits of accuracy. If we want the error probability to be δ we need the error probability to be $\frac{\delta}{k}$ per bit. Therefore, the algorithm will require $O(klog\frac{k}{\delta})$ measurements.
We would like to express this algorithm as a quantum unitary transformation. First we will use the first algorithm with $U, U^2, ...$ We can make sure that $|e\rangle$ will remain after each operation, and therefore we will be able to use it for the next operation. We have a classic algorithm for extrapolating θ from the estimations of $\theta, 2\theta, 4\theta, ...$ which works with high probability δ. Hence, there is a quantum computation which computes $|e\rangle \rightarrow |e\rangle(a|\theta\rangle + b|\theta'\rangle)$ where θ is the correct answer and $a >> b$.
If U satisfies $U^S = I$, Then for each $j, \lambda_j = e^{2\pi i\theta}$ where $\theta = \frac{a}{b}$ for integers a, b satisfying $b \leq s$. Thus we can recover the precise value of λ by finding θ to within $\frac{1}{s^2}$ and then using continued fractions to find the closest such ratio. Hence we need only $log(s^2)$ bits of θ.

1.3 Kitaev's Lemma

Given an unitary matrix U, and a polynomial algorithm to compute U^{2^l}, Then it is possible to efficiently compute the unitary transformation :
$|e\rangle|0\rangle \rightarrow |e\rangle|\theta\rangle$. Whereas $\lambda = e^{2\pi i\theta}$ is the corresponding eigenvalue of the eigenvector $|e\rangle$ of U.

2 Hidden Subgroup Problem

Given a finite group G, and an oracle with computes a function $f : G \rightarrow \{0,1\}^n$ which is constant on every coset of $H \preceq G$,
i.e. $\forall g \in G\ \exists a$ s.t. $\forall h \in H\ f(gh) = a$.
The Hidden subgroup problem is to find H.

2.1 Versions of HSP

Many of the problems we have seen in previous lessons are HSP. for example

1. **Simon** : The goal is to find the cycle of a cyclic function f. This is a HSP where $G = {Z_2}^n, H = \{0, a\}$. f is constant on the cosets of H

because $f(x) = f(x \oplus a)$, and when we will find H, we will have the cycle a.

2. **Deutsch – Josza** : Given $f_a : x \to x \cdot a$, find a. This is a HSP where $G = {Z_2}^n, H = \{x | x \perp a\}$. f is constant on the cosets of H because it is unitary (i.e. preserves angles), and so vectors which where perpendicular remain so, and vise versa. Knowing H we can find a, in a similar way to that we used in Simon's algorithm.

3. **Shor** : The goal is to find the order of $a \in Z_Q$. If $r|Q$ then $G = {Z_2}^n, H = \{0, r, 2r, 3r, ...\}$. f is constant on the cosets of H, and from H we can salvage r and from r we can salvage a. If r doesn't divide Q, it's not a HSP, but it's similar.

4. **Graph Automorphism** : Given a graph K, the goal is to find

   $Aut(K) = \{\sigma \in S_n | \sigma(K) = K\}$. This is a HSP where $G = S_n$, $H = Aut(K)$ and $f : \sigma \to \sigma(K)$. If we can solve the GA problem, we can also solve the Graph Isomorphism problem because

   $$K1 \cong K2 \Leftrightarrow |Aut(K1 \cup K2)| = 2|Aut(K1)| = 2|Aut(K2)|$$

   .

3 Representations of finite groups

Definition 1 *A k-dimension* **Representation** *of the group G is a mapping* $\rho : G \to \{M_k(\mathbb{C})\}$ *such that* ρ *is a homomorphism of G, i.e.* $\rho(gh) = \rho(g)\rho(h)$.

Definition 2 *A group representation* ρ *is* **reducible** *if* $\rho(G) = V_1 \oplus V_2$ *and* $\rho(g)$ *operates on* V_1 *and* V_2 *separately (meaning* $\rho(g)$ *is a block matrix). Therefore we can regard* ρ *as two representations over* V_1 *and* V_2.

Definition 3 *A group G is called* **Abelian** $\Leftrightarrow \forall g_1, g_2 \in G \; g_1 g_2 = g_2 g_1$

Representations of Abelian groups

The representations of Abelian groups are commutative and therefore simultaneously diagonal. Hence, all the irreducible representations are one-dimensional. Lets look at all the representations of the cyclic group Z_Q. if ρ is not trivial (i.e. constant) then

$$\rho(g) = \rho(g+0) = \rho(g)\rho(0)$$

$$\Downarrow$$

$$\rho(0) = I$$

furthermore

$$(\rho(1))^Q = \rho(1 + \cdots + 1) = \rho(0) = 1$$

$$\Downarrow$$

$$\rho(1) = {\omega_Q}^j = e^{\frac{2\pi ij}{Q}}$$

$\rho(1)$ determines ρ since $\rho(i) = (\rho(1))^i = {\omega_Q}^{ij}$. Therefore we have found all the Q possible one-dimensional representations (which are also called characters) The rest of the representations can be written as a direct sum of the characters. The set of characters is an orthonormal basis of the function space over G.

Theorem 1 *(The structure theorem of Abelian groups.) If G is Abelian then* $G \cong Z_{N_1} \times \cdots \times Z_{N_k}$

Given an Abelian group G, its representation, ρ, satisfies

$$\rho(g) = \rho(g_1 \cdots g_k) = \rho(g_1) \cdots \rho(g_k) = {\omega_{N_1}}^{j_1 g_1} \cdots {\omega_{N_k}}^{j_k g_k}$$

Definition 4 $\hat{\mathbf{G}}$ *is the set of all non-separable representations of G*

If G is Abelian then $G \cong \hat{G}$. We have seen that for cyclic groups, and the theorem above expands that notion to all Abelian groups.

4 Fourier Transforms over a finite group

It can be easily checked that the set of irreducible representations of an Abelian group is an orthonormal basis to the Hilbert space of functions from G to the complex numbers. (In fact, we have seen this several times before.)

We will not prove it here, but a generalization of this is true for non-Abelian groups: the vectors $|\alpha(\rho,i,j)\rangle = (\rho_{i,j}(g_1), \rho_{i,j}(g_2), ...\rho_{i,j}(g_{|G|}))$, for all ρ irreducible representation of G and all indices i, j also form an orthonormal basis for the functions from the group G to the complex numbers. (In other words,

$$\langle \alpha(\rho,i,j)|\alpha(\rho',i',j')\rangle = 0$$

unless $\rho = \rho', i = i', j = j'$.

The Fourier transform is defined to be the transformation to this orthonormal basis. For Abelian groups, it is defined as follows:

$$|g\rangle \rightarrow \frac{1}{\sqrt{|G|}} \sum_{\rho \in \hat{G}} \rho(g)|\rho\rangle$$

We give the definition of Fourier transform in the general case, even though we will not use it here. If d_ρ is the dimension of the representation ρ, we define the Fourier transform to be:

Fourier Transform of a group G is defined by

$$|g\rangle \rightarrow \sum_{\rho \in \hat{G}, i,j \leq d_\rho} \sqrt{\frac{d_\rho}{|G|}} \rho_{ij}(g)|\rho, i, j\rangle$$

where $|\rho, i, j\rangle$ are simply computational basis states running over the names and indices of all irreducible representations of G.

4.1 Kitaev's algorithm for FT over Z_Q

We want to compute :

$$|\psi_{Q,a}\rangle = \frac{1}{\sqrt{Q}} \sum_{b=0}^{Q-1} e^{\frac{2\pi i a b}{Q}} |b\rangle$$

Kitaev suggested the following algorithm :

$$|a\rangle \otimes |0\rangle$$

$$\downarrow 1$$

$$|a\rangle \otimes |\psi_{Q,0}\rangle$$

$$\downarrow 2$$

$$|a\rangle \otimes |\psi_{Q,a}\rangle$$

$$\downarrow 3$$

$$|0\rangle \otimes |\psi_{Q,a}\rangle$$

$$\downarrow 4$$

$$|\psi_{Q,a}\rangle \otimes |0\rangle$$

Lets explain the transitions:

1. $|\psi_{Q,0}\rangle = \frac{1}{\sqrt{Q}}\sum|i\rangle$, and therefore this transition is equivalent to picking an integer between 1 and Q uniformly at random. We will do so using a recursive procedure. Let $2^{n-1} < Q < 2^n$. Denote $Q_0 = 2^{n-1}$ and $Q_1 = Q - Q_0$. Apply the one qubit gate $(|0\rangle \to \sqrt{\frac{Q_0}{Q}}|0\rangle + \sqrt{\frac{Q_1}{Q}}|1\rangle)$. Now, conditioned on the first bit x, create on the last $n-1$ bits, the state $|\psi_{Q_x,0}\rangle$ recursively.

2. $|a\rangle \otimes |\psi_{Q,0}\rangle \to |a\rangle \otimes |\psi_{Q,a}\rangle$ is achieved by applying $|a,b\rangle \to e^{\frac{2\pi iab}{Q}}|a,b\rangle$.

3. To apply the third step we shall first prove a small lemma:

   Lemma 2 *Let* $U = \begin{pmatrix} 0 & 1 & & & 0 \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ & & & \cdot & 1 \\ 1 & 0 & & & 0 \end{pmatrix}$ *, Then* $U|\psi_{Q,a}\rangle = e^{\frac{-2\pi ia}{Q}}|\psi_{Q,a}\rangle$

   Proof : Notice that $U|i\rangle = |i+1\rangle$ and so :

   $$U|\psi_{Q,a}\rangle = \sum e^{\frac{2\pi ia(b+1-1)}{Q}}|b+1\rangle = e^{\frac{-2\pi ia}{Q}}|\psi_{Q,a}\rangle$$

   QED.

   Therefore transition 3 is actually the reverse of computing the eigenvalue of the eigenvector $|\psi_{Q,a}\rangle$, which is, by Kitaev's Lemma, a possible transition.

4. The last transition is merely swapping the bits.

5 H and $H^{\perp}$.

Definition 5 $\mathbf{H}^{\perp} = \{\rho \in G | \forall h \in H \;\; \rho(h) = 1\}$

Lemma 3 $|gH\rangle \rightarrow^{FT} \frac{\sqrt{|H|}}{\sqrt{|G|}} \sum_{\rho \in H^{\perp}} \rho(g) |\rho\rangle$

Proof

$$|H\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |h\rangle \rightarrow^{FT} \frac{1}{\sqrt{|H|}} \sum_{h \in H} [\frac{1}{\sqrt{|G|}} \sum_{\rho \in \hat{G}} \rho(h) |\rho\rangle] =$$

$$\frac{1}{\sqrt{|H||G|}} \sum_{\rho \in \hat{G}} [\sum_{h \in H} \rho(h)] |\rho\rangle$$

if $\rho \in H^{\perp}$ then $\sum \rho(h) = |H|$, because $\forall h \in H \;\; \rho(h) = 1$. Otherwise, there is h' such that $\rho(h') \neq 1$ and so

$$\sum \rho(h) =^{\clubsuit} \sum \rho(hh') = (\sum \rho(h)) \rho(h')$$

$$\Downarrow$$

$$\sum \rho(h) = 0$$

$\clubsuit$ since we sum over all $h \in H$ it is equivalent to sum over all hh' for some constant h'. QED.
Therefore we have :

$$|H\rangle \rightarrow^{FT} \frac{1}{\sqrt{|H||G|}} \sum_{\rho \in H^{\perp}} |H| |\rho\rangle = \frac{\sqrt{|H|}}{\sqrt{|G|}} \sum_{\rho \in H^{\perp}} |\rho\rangle$$

In a similar way we can prove :

$$|gH\rangle \rightarrow^{FT} \frac{\sqrt{|H|}}{\sqrt{|G|}} \sum_{\rho \in H^{\perp}} \rho(g) |\rho\rangle$$

5.1 Algorithm for HSP over Abelian groups

Finally, we can describe the algorithm which finds H (!), using the oracle f :

$$|0\rangle \rightarrow^{FT} \frac{1}{\sqrt{|\hat{G}|}} \sum_{g\in G} \rho(0)|g\rangle = \frac{1}{\sqrt{|\hat{G}|}} \sum_{g\in G} |g\rangle \rightarrow^{f} \frac{1}{\sqrt{|\hat{G}|}} \sum_{g\in G} |g\rangle|f(g)\rangle$$

Now if we will measure $|f(g)\rangle$ we will get $|gH\rangle$ since f is constant on $|gH\rangle$ and different on other cosets. As we seen in the previous lemma, we can use Fourier Transform to get $\frac{\sqrt{|H|}}{\sqrt{|G|}} \sum_{\rho\in H^{\perp}} \rho(g)|\rho\rangle$. Now by measuring the result we get a random member of $H^{\perp}$, $\rho_{j_1...j_k}$ (The distribution is uniform because $|\rho(g)| = 1$). Notice that :

$$\rho_{j_1...j_k}(g_1...g_k) = {\omega_{N_1}}^{j_1 g_1} \cdots {\omega_{N_k}}^{j_k g_k} = {\omega_{N'}}^{q_1 j_1 g_1} \cdots {\omega_{N'}}^{q_k j_k g_k} = {\omega_{N'}}^{\sum q_i j_i g_i}$$

where: $N' = lcm(N_1, ..., N_k), q_i = \frac{N'}{N_i}$, and therefore $\omega_{N_i} = {\omega_{N'}}^{q_i}$.
$\rho_{j_1...j_k} \in H^{\perp} \Rightarrow \rho_{j_1...j_k}(g_1...g_k) = 1$, and so the generalized inner product $\sum q_i j_i g_i = 0 \ (modN')$. Remember that each ρ defines $j_1...j_k$ and vice versa. We can recover H, in a way similar to Simon's algorithm : If we measure enough ρ's, they will span $H^{\perp}$. We will show that the number of ρ's that we need is polynomial. Let n donate the dimension of $H^{\perp}$. Assume we have spanned a k-dimensional linear space, we can represent it as the first k coordinates of an n-dimensional vector. Therefore, the probability to measure a vector in that space is the same as getting 0 out of $[0,..,N']$ n-k times, or $(\frac{1}{N'})^{n-k}$. for $k = n-1$ the probability is $\frac{1}{N'}$. Therefore the probability that a new ρ will be linearly dependent in the other ρ's is $O(\frac{1}{N'})$.